An approximate dynamic programming approach for sequential pig marketing decisions at herd level*

Reza Pourmoayed† and Lars Relund Nielsen

Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark.

February, 2017

Abstract: One of the most important operations in the production of growing/finishing pigs is the marketing of pigs for slaughter. While pork production can be managed at different levels (animal, pen, section, or herd), it is beneficial to consider the herd level when determining the optimal marketing policy due to inter-dependencies, such as those created by fixed transportation costs and cross-level constraints. In this paper, we consider sequential marketing decisions at herd level. A high-dimensional infinite-horizon Markov decision process (MDP) is formulated which, due to the curse of dimensionality, cannot be solved using standard MDP optimization techniques. Instead, approximate dynamic programming (ADP) is applied to solve the model and find the best marketing policy at herd level. Under the total expected discounted reward criterion, the proposed ADP approach is first compared with a standard solution algorithm for solving an MDP at pen level to show the accuracy of the solution procedure. Next, numerical experiments at herd level are given to confirm how the marketing policy adapts itself to varying costs (e.g., transportation cost) and cross-level constraints. Finally, a sensitivity analysis for some parameters in the model is conducted and the marketing policy found by ADP is compared with other well-known marketing polices, often applied at herd level.

Keywords: OR in agriculture; approximate dynamic programming; Markov decision process; herd management; stochastic dynamic programming.

1 Introduction

One of the most important operations in the production of growing/finishing pigs is the marketing (culling) of pigs for slaughter. Each week, the farm manager must decide which pigs should be delivered to the abattoir and when the pen (or section) should be emptied (Kure 1997). In the production system, animals may be considered at different levels: herd, section, pen, or animal. The herd is a group of sections, a section includes a number of pens, and a pen involves some animals (usually

*Preprint of R. Pourmoayed and L.R. Nielsen, *An approximate dynamic programming approach for sequential pig marketing decisions at herd level* in European Journal of Operational Research (2019), doi:10.1016/j.ejor.2019.01.050

†Corresponding author, email: rpourmoayed@eng.au.dk.

15-20 animals). Marketing decisions can be considered at different levels, e.g., animal (Glen 1983) or pen (Kristensen, Nielsen, and Nielsen 2012) and (Kure 1997). The complexity of the marketing decisions depends on the number of levels that are taken into account simultaneously and on how decisions at different levels are linked together. In other words, cross-level constraints between pen, section, and herd level can affect marketing decisions and they should be considered in the problem. For example, all pens in a section have to be terminated at the same time and it may be optimal to terminate a section before all the remaining pigs in the section reach their optimal individual weight to keep the length of a production cycle short. In this paper, we focus on marketing decisions at herd level. In the following, we describe the problem in more detail.

A *production cycle* in the production process of finishing pigs in a section is started by buying piglets on the market or transferring them from another production unit when they weigh approximately 30 kilograms. In the section of the finisher herd, the piglets grow until marketing (9-15 weeks). Because pigs grow at different growth rates, they obtain their slaughter weight at different times in the last weeks of the growing period. Hence, at the end of the growing period, the decision maker should determine which pigs should be selected for slaughter in each pen (*individual marketing*) and after a sequence of individual marketings, the decision maker must decide when to *terminate* (empty) each section. Terminating a section means that the remaining pigs in the section are sent to the abattoir (in one delivery) and after cleaning the section, another group of piglets (each weighing approximately 30 kilograms) is inserted into the pens of the corresponding section and the production cycle is repeated. During the marketing period, marketed pigs from the pens are grouped in one weekly delivery and are transported to the abattoir by means of a number of trucks. Depending on the number of marketed pigs at herd level and the capacity of each truck, the decision maker should also determine the number of trucks needed to transport the pigs to the abattoir. That is, transportation costs may have an effect on the marketing policy of the fattening pigs. In most Danish herds, transportation of culled pigs is handled by a single abattoir and the transportation cost is fixed per truckload.

The reward of marketing depends on the price of the carcass weight in the abattoir, the cost of buying the piglets, the cost of feeding the pigs, and the cost of transporting the culled pigs to the abattoir. Other costs such as labour and veterinary services are assumed fixed and are not included in the model. The best meat price is obtained if the carcass weight lies within a specific interval. Therefore, the farmer must time the marketing decisions while simultaneously considering the carcass weight in relation to the best interval, the transportation cost of trucks (more pigs per truck better utilizes capacity), and the length of the production cycle for feeding the rest of the pigs (small production cycles give a higher throughput per time unit). For an extended overview over pig production of growing/finishing pigs, see Pourmoayed and Nielsen (2014).

Marketing decisions have been studied by a number of researchers. Kure (1997) proposed a recursive dynamic programming method and used replacement theory concepts to find the best marketing strategy. In the study by Jørgensen (1993), a *hierarchical Markov decision process (HMDP)* was applied to analyse the precision of the weighing methods on the marketing policy of fattening pigs. Toft, Kristensen, and Jørgensen (2005) combined decisions related to the delivery strategy of pigs to the abattoir and epidemic diseases using an HMDP. In the study by Kristensen et al. (2012), an HMDP was employed to model marketing decisions with online weight information and the state space of the HMDP was related to state space models using Bayesian updating. In the study by Plà-Aragonés, Rodriguez-Sanchez, and Rebillas-Loredo (2013), the optimal marketing policy was found by a mixed integer linear programming method under an all-in all-out strategy. They formulated the problem by a mathematical programming model and solved their model using a heuristic approach under different pig size distributions and pig growth rates. Niemi (2006) applied a stochastic dynamic programming method to find the best time of marketing for an individual pig and the best nutrient ingredients in the feed-mix, simultaneously. In a recent study, Pourmoayed, Nielsen, and Kristensen (2016) have considered optimal marketing and feeding strategies at pen level.

To the best of our knowledge, few studies consider the problem at herd level and take the effect of transportation costs on marketing decisions into account. Boys, Li, Preckel, Schinckel, and Foster (2007) analysed the effect of single and multiple shipping decisions on the marketing strategy of a heterogeneous herd using a simulation method. For each simulation instance they optimize the shipment times to the abattoir. No fixed transportation cost is assumed and the pigs must be shipped in full truckloads if possible. Ohlmann and Jones (2011) used a mixed integer programming (MIP) model to find the best marketing strategy in a barn of pigs. Pig growth is modelled using a set of weight classes and the weekly growth is modelled using growth transition probabilities. A fixed transportation cost is assumed. A sequence of MIP models are solved for different length of the production cycle to identify the cycle length that maximizes the objective. In both studies, a single section is considered (the whole barn must be cleaned) and marketing decisions are examined under an annual profit criterion.

In this paper, we consider marketing decisions in a herd composed of sections and pens. We formulate the sequential marketing decisions using a discounted infinite-horizon *Markov decision process (MDP)* and assume that the production process is cyclic at section level, i.e., when a section is terminated, a new batch of piglets is inserted into the pens of this section and a new production cycle is started. The model is stochastic due to the uncertainty of the growth of the pigs. This uncertainty is described by a stochastic process relying on *state space models* formulated in Pourmoayed et al. (2016). Due to the large number of states and actions in the model, the *curse of dimensionality* becomes apparent and the usual solution procedures for MDPs (e.g., policy iteration) can not be used to solve the model. Therefore, we use an approximation strategy to identify a marketing policy at herd level. More precisely, we first use the properties of the value function in a discounted infinite-horizon MDP at pen level to find an approximation architecture for the value function at herd level, and next we apply an *approximate dynamic programming (ADP)* approach with *post-decision states* to find the best marketing policy at herd level (Powell 2007).

The main idea of approximate dynamic programming is to approximate the value function by finding the best actions for a set of states observed in the system. The approximation architecture of the value function is often described by a parametric function and there are well-known algorithms exploiting simulation and linear programming techniques to estimate the parameters (see e.g., de Farias and Van Roy 2003, Powell 2010, Topaloglu and Powell 2006). For more details about ADP algorithms, the interested reader may refer to Powell (2007). Examples of approximation strategies used to optimize livestock systems are Kristensen (1992) and Ben-Ari and Gal (1986) that exploit a parameter iteration algorithm developed by Gal (1989) to find the best replacement policy in a dairy herd. In these studies, however, the estimation of transition probabilities and the calculation of expected value operators in the solution procedure are computationally challenging and require much computational effort, due to complicated transition functions and large state and action spaces. In the present study, we have reformulated the Bellman equations in the form of *post-decision states* and we thereby gain a significant improvement in the computational efficiency (CPU time) of the solution procedure and present a novel approach for optimizing livestock systems modelled by high-dimensional MDPs. To summarize, the main contributions of this paper are as follows.

We formulate a novel MDP model for sequential decision-making at herd level optimizing the total expected discounted reward over an infinite time horizon. The model is a high-dimensional MDP and differs from the model in Pourmoayed et al. (2016) which only consider marketing decisions at pen level (modelled as an HMDP). The MDP model differs from previous papers focusing on herd level (Boys et al. 2007, Ohlmann and Jones 2011) in both methodology and assumptions. In both Boys et al. (2007) and Ohlmann and Jones (2011) a single section is assumed. That is, the whole barn must be terminated and cleaned before new piglets can be inserted. We extend this to multiple sections, that may be terminated at different times depending on the production state in each section.

Our model optimizes decisions dynamically. That is, decisions are optimized based on the current state of the system, the present reward, and the expected value of the future states in the system. As a result, the length of the production cycle under a policy is stochastic and different decisions may be taken depending on the realization of states. The uncertainty is described by stochastic processes relying on state space models formulated in Pourmoayed et al. (2016). That is, a Bayesian approach is used to update the system such that it contains the relevant information based on the realized observations. The state space models are embedded into the MDP such that the state of each pen is found using the current data stream for each pen in the herd. That is, the model takes into account the impact of the sectioning of the production facility. As a result, the model has an operational focus.

Because our MDP suffers from the curse of dimensionality, we formulate a tailored ADP algorithm to find the best marketing policy at herd level which is validated on a small sized model. Extensive computational results are given for a standard Danish herd. First, we analyse the effect of increasing fixed transportation cost on marketing decisions. Next, we show that the policy found using ADP outperforms other well-known marketing policies in the industry. Finally, we examine the effect of changes in feed cost and pork price.

The rest of the paper is organized as follows. In Section 2, sequential marketing decisions are modelled using a discounted infinite-horizon MDP model. Section 3 describes an approximate dynamic programming approach for solving the MDP model. In Section 4, numerical examples are given and the policy resulting from the ADP is compared with other marketing policies, and finally in Section 5, conclusions are given.

2 Model description

An *infinite-horizon MDP* models a sequential decision problem over an infinite time horizon. We assume weekly decision epochs and let $\mathbb{S}$ denote the finite set of system states at an arbitrary decision epoch. Given system *state* $s \in \mathbb{S}$ at the current *stage* or decision epoch, an *action* a from the finite set of allowable actions $\mathbb{A}(s)$ is chosen resulting in an immediate *reward* $r(s,a)$ and a probabilistic transition to state $s_+ \in \mathbb{S}$ at the next decision epoch. This transition is based on the *transition function* $\phi(s,a,\omega)$ where ω denotes random information received between the current and next decision epochs. Random information ω depends on a stochastic process affecting the system state. A *policy* π is a decision rule or function that assigns for each state an action $a = \pi(s) \in \mathbb{A}(s)$, i.e., a policy prescribes which action to take for each state. Our goal is to find a policy that maximizes the *total expected discounted reward* over an infinite time horizon:

$$\max_{\pi} \mathbb{E}\left(\sum_{n=1}^{\infty} \gamma^{n-1} r_n(s, \pi(s)) \right), \tag{1}$$

where γ is the discount factor used to calculate the net present value and subscript n is added to show the stage number. We assume that $\mathbb{S}$ and $\mathbb{A}(s)$ are finite, $0 < \gamma < 1$, and $r(s,a)$ is bounded. Hence an optimal policy can be found among the set of deterministic stationary policies (Puterman 2005, Theorem 6.2.10) and a unique solution to the following optimality equations will exist (Puterman 2005, Section 6.2):

$$\nu(s) = \max_{a \in \mathbb{A}(s)} \left(r(s,a) + \gamma \mathbb{E}\left(\nu(s_+)\right) \right), \quad \forall s \in \mathbb{S}, \tag{2}$$

where $s_+ = \phi(s,a,\omega)$, and the value function $\nu(s)$ denotes the expected discounted reward of being in state s.

We consider the following assumptions when modelling sequential marketing decisions at herd level:

1. A herd consists of $|I|$ sections and each section $i \in I$ includes $|J|$ pens (for simplicity and without loss of generality, we assume that the number of pens is the same in all sections).

2. In the beginning of a production cycle in a section, each pen is filled with $q^{\max}$ pigs.

3. The marketing decisions are made on a weekly basis and the culled pigs are transferred to the abattoir after a few days.

4. Individual marketing at pen level is started in week $t^{\min}$ at the earliest.

5. A section is terminated in week $t^{\max}$ at the latest, i.e., the maximum life time of a pig is $t^{\max}$ weeks.

6. Weekly deliveries to the abattoir are based on a cooperative agreement where culled pigs from each section in the herd are grouped and transferred to the abattoir by trucks. Transportation of culled pigs is handled by the abattoir and the transportation cost is fixed per truck. Hence, variable costs of transportation, e.g., costs of loading a pig into the truck, are not considered in the model.

7. A section is cleaned after termination. After the cleaning period, a new batch of piglets is inserted into the pens immediately, i.e., piglets are always available. Moreover, information about the current growth and weight of the piglets is known at insertion time.

8. The production process is independent among sections, i.e., piglets can be inserted into different sections at different times, and a section can be terminated independent of other sections.

9. The sequence of feed-mixes used during the production cycle (feeding strategy) is known and fixed.

10. A new batch of piglets and the required feed stock are bought at known and fixed prices.

11. The pigs are sold to the abattoir according to a known settlement pork price function.

To define the MDP, we describe the state space, action space, rewards, and transition function in the following subsections.

2.1 State space

In a given decision epoch, state $s \in \mathbb{S}$ is defined using state variables:

t_i week number in a production cycle of section $i \in I$;

q_{ij} number of remaining pigs in pen $j \in J$ of section $i \in I$;

$w_{ij} = (\mu_{ij}, \sigma_{ij}, g_{ij})$ weight information of the pigs in pen $j \in J$ of section $i \in I$. The three state variables μ_{ij}, σ_{ij}, and g_{ij} correspond to the estimated mean and standard deviation of the pig weights and the average growth of the pigs in the pen, respectively. The values of these variables can be obtained by using repeated measurements of weight and feed intake data at pen level applied to a state space model using Bayesian updating (see Pourmoayed et al. (2016)).

That is, state s is represented as $s = (t, q, w)$ where

$$\begin{aligned} t &= (t_1, .., t_i, .., t_{|I|}), \\ q &= (q_{11}, .., q_{ij}, .., q_{|I||J|}), \\ w &= (w_{11}, .., w_{ij}, .., w_{|I||J|}), \end{aligned}$$

and the set of states becomes

$$\mathbb{S} = \{s = (t, q, w) \mid t_i \in \{1, \dots, t^{\max}\},$$
$$q_{ij} \in \{0 \cdot \mathbf{I}_{\{t_i > t^{\min}\}} + q^{\max} \mathbf{I}_{\{t_i \leq t^{\min}\}}, \dots, q^{\max}\}, w_{ij} \in \mathbb{W}, i \in I, j \in J\},$$

where $\mathbf{I}_{\{\cdot\}}$ denotes the indicator function. The set $\mathbb{W}$ denotes the set of weight information that is the result of the discretisation of the possible outcomes for the statistical models embedded into the MDP; this discretisation is described in Section 4.1. The statistical models are taken from Pourmoayed et al. (2016) and a short review is given in Appendix A.

2.2 Action space

Assume that the pigs are sorted in ascending order based on their live weight in each pen such that index k denotes the k^{th} lightest pig in the pen. We consider the following decision variables for defining action a:

x_{ijk} equals 1 if the k^{th} pig $(k \in \{1, 2, \dots, q_{ij}\})$ in pen $j \in J$ of section $i \in I$ is culled and zero otherwise;

y_i equals 1 if section $i \in I$ is terminated and zero otherwise.

Hence action a is defined as $a = (x, y)$ where

$$x = (x_{111}, .., x_{ijk}, .., x_{|I||J|q_{|I||J|}}),$$
$$y = (y_1, .., y_j, .., y_{|I|}),$$

and the set of possible actions $\mathbb{A}(s)$ is defined using the following constraints:

$$x_{ijk} \leq x_{ijk+1}, \qquad \forall i \in I, j \in J, k \in \{1, 2, \dots, q_{ij} - 1\}, \tag{3}$$
$$\sum_{j \in J} q_{ij} - \sum_{j \in J} \sum_{k=1}^{q_{ij}} x_{ijk} \leq M(1 - y_i), \quad \forall i \in I, \tag{4}$$
$$\sum_{j \in J} q_{ij} - \sum_{j \in J} \sum_{k=1}^{q_{ij}} x_{ijk} \geq (1 - y_i), \qquad \forall i \in I, \tag{5}$$
$$\sum_{i \in I} \sum_{j \in J} \sum_{k=1}^{q_{ij}} x_{ijk} \leq k^{\texttt{truck}} z, \tag{6}$$
$$z \text{ integer}, \tag{7}$$
$$x_{ijk} \text{ binary}, \qquad \forall i \in I, j \in J, k \in \{1, 2, \dots, q_{ij}\}, \tag{8}$$
$$y_i \text{ binary}, \qquad \forall i \in I. \tag{9}$$

Due to the order of the sorted pigs in the pen, constraint (3) enforces that a lighter pig cannot be culled earlier than the heavier pigs. If section termination happens ($y_i = 1$), big-M constraint (4) implies that all the remaining pigs in the pens of section i must be culled. Moreover, if all pigs are culled in section i, (5) ensures that this section will be terminated ($y_i = 1$). Constraint (6) expresses that the total number of marketed pigs in the herd must be less than the capacity of the trucks called from the abattoir. In this constraint, $k^{\texttt{truck}}$ is the capacity of one truck and z is a supplementary integer variable denoting the number of trucks needed to transport the marketed pigs to the abattoir which is uniquely determined given action a.

In addition, if state $s = (t, q, w)$ is observed such that $1 \leq t_i \leq t^{min}$ then $x_{ijk} = 0$ for all $j \in J$ and $k \in 1, 2, ..., q_{ij}$ because marketing is not allowed. Also if $t_i = t^{max}$, then $y_i = 1$ since the section must be terminated in the last week of the production cycle.

2.3 Rewards

The reward of action $a \in \mathbb{A}(s)$ is calculated as the revenue of selling the marketed pigs to the abattoir minus the cost of feeding the remaining pigs until the next decision epoch, the cost of transferring the marketed pigs to the abattoir by trucks, and the cost of buying a new batch of piglets (if a section is terminated). Hence, the reward associated with state s and action a is formulated as

$$r(s,a) = \sum_{i \in I} \sum_{j \in J} \sum_{k=1}^{q_{ij}} c_k^{\texttt{cull}}(w_{ij}) x_{ijk} - \sum_{i \in I} \sum_{j \in J} \sum_{k=1}^{q_{ij}} c_k^{\texttt{feed}}(w_{ij})(1 - x_{ijk}) - c^{\texttt{truck}} z - \sum_{i \in I} c^{\texttt{piglet}} |J| q^{\max} y_i, \tag{10}$$

where $c_k^{\texttt{cull}}(w_{ij})$ is the expected unit reward of selling the k^{th} pig in pen j of section i to the abattoir given weight information w_{ij}. Similarly, $c_k^{\texttt{feed}}(w_{ij})$ denotes the expected feeding cost of the k^{th} pig in pen j of section i kept in the herd until the next decision epoch. Note that when marketing decisions are made, culled pigs are sent to the abattoir after a few days and therefore the additional feeding cost and reward, resulting from the weight gain of culled pigs in this period, are considered in the calculation of $c_k^{\texttt{cull}}(w_{ij})$. We use a simulation method to calculate functions $c_k^{\texttt{cull}}(w_{ij})$ and $c_k^{\texttt{feed}}(w_{ij})$ given weight information w_{ij} (for more details, see Appendix B). The coefficient $c^{\texttt{truck}}$ is the fixed truck cost for transporting the culled pigs to the abattoir, and $c^{\texttt{piglet}}$ is the cost of buying a new piglet in the beginning of a production cycle.

2.4 Transition function

Given state $s = (t, q, w)$ and action a, the transition function $\phi(s, a, \omega) = (t_+, q_+, w_+) = s_+$ describes how the system evolves from state s to state s_+. That is, for every $i \in I$ and $j \in J$, we have

$$\begin{aligned}
t_{i_+} &= (1 - y_i)(t_i + 1) + y_i, \\
q_{ij_+} &= (1 - y_i)(q_{ij} - \sum_{k=1}^{q_{ij}} x_{ijk}) + y_i q^{\max}, \\
w_{ij_+} &= (1 - y_i)\Gamma(w_{ij}, \omega) + y_i w_0,
\end{aligned}$$

where w_0 denotes the weight information at the start of a production cycle and the stochastic function $\Gamma(w_{ij}, \omega)$ describes a stochastic transition between w_{ij} and w_{ij_+} given random information ω. This transition relies on two state space models using Bayesian updating presented in Pourmoayed et al. (2016). Hence, given weight information $w_{ij} = (\mu_{ij}, \sigma_{ij}, g_{ij})$, in state s, the transition probabilities can be calculated. For more details about this stochastic process, see Appendix A.

3 Approximate dynamic programming

Solving the model using the optimality equations in (2) is not possible in practice. First, the sizes of the state and action spaces in the model are prohibitive and hence computation of $\nu(s)$ for every possible state is impossible. Second, due to the large number of states and possible outcomes for random information ω, an exact computation of the expected value $\mathbb{E}\left(\nu(s_+)\right)$ in (2) is expensive. Finally, due to the expected value operator, the maximization problem in (2) is not deterministic and hence it may be difficult to solve it and find the optimal actions. These computational challenges are known as the three curses of dimensionality (Powell 2007, Section 4.1) that prevent us from applying standard solution algorithms (e.g., policy and value iteration) to solve the model.

In this section, we formulate an approximate dynamic programming (ADP) algorithm using post-decision states to find an approximate solution for the model. First, in Section 3.1, we design a parametric approximation architecture for the value function $\nu(s)$. Second, in Section 3.2, the optimality equations are reformulated in terms of post-decision states to find a deterministic version of the maximization problem in (2). Finally, in order to estimate the parameters of the approximated value function, an approximate value iteration algorithm is employed and presented in Section 3.3.

3.1 Approximation architecture of the value function

In order to find an approximation architecture for the value function, we have to consider a balance between the computational efficiency and the quality of the resulting policy. For instance, a second or third polynomial approximation for the value function $\nu(s)$ may change the maximization problem in (2) to a very difficult non-linear model (Powell and Van Roy 2012). In this study, we use our knowledge about the value function at pen level to find an approximation architecture for the value function at herd level (Pourmoayed et al. 2016).

Due to the hierarchical structure of the herd with sections and pens, we assume that the value function at herd level is additive over the value functions at pen level. That is, an approximation of the value function is:

$$\nu(s) \approx \sum_{i \in I} \sum_{j \in J} \nu_{ij}(t_i, q_{ij}, w_{ij}), \tag{11}$$

where $\nu_{ij}(t_i, q_{ij}, w_{ij})$ denotes the value function related to pen j of section i including q_{ij} pigs in week t_i with weight information w_{ij}. According to the form of the value function when optimizing at pen

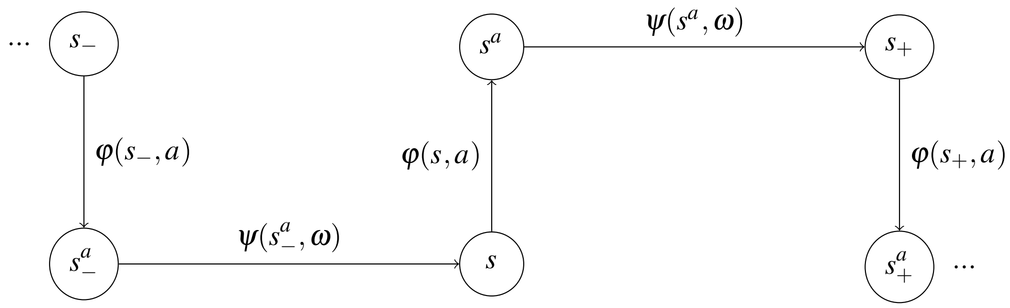

Figure 1: The relation between pre- and post-decision states over three stages. $\varphi(\cdot)$ indicates a deterministic transition from a pre-decision state to a post-decision state, and $\psi(\cdot)$ shows a stochastic transition from a post-decision state to a pre-decision state. Subscripts '-' and '+' clarify a state in the previous and the next decision epochs, respectively.

level (Pourmoayed et al. 2016), a good functional form of $v_{ij}(t_i, q_{ij}, w_{ij})$ is

$$v_{ij}(t_i, q_{ij}, w_{ij}) \approx q_{ij} b(t_i, w_{ij}), \tag{12}$$

where $b(t_i, w_{ij})$ is a parameter describing the effect of weight information w_{ij} in week t_i at pen level. Note that $b(t_i, w_{ij})$ is the value of having exactly one pig more in the pen and hence it is the gradient or *slope parameter* of the value function with respect to the number of remaining pigs, q_{ij}.

By substituting (12) in (11), the final form of the approximation architecture for the value function at herd level becomes

$$v(s) \approx \sum_{i \in I} \sum_{j \in J} q_{ij} b(t_i, w_{ij}), \tag{13}$$

which eliminates the need for updating the value function for every possible state as we now only need to estimate the slope parameter $b(t_i, w_{ij})$ for $t_i \in \{1, ..., t^{\max}\}$ and $w_{ij} \in \mathbb{W}$. Note that since the number of slope parameters is much lower than the number of possible states, the computational effort of the solution algorithm will be noticeably reduced.

3.2 Post-decision states

The transition function $\phi(s, a, \omega)$ can be split such that the effect of action a and random information ω is separated. For a given state $s = (t, q, w)$ and action a, the transition to $s_+ = \phi(s, a, \omega)$ is divided into two steps:

$$\begin{aligned} s^a &= \varphi(s, a) = (t^a, q^a, w^a), \\ s_+ &= \psi(s^a, \omega) = (t_+, q_+, w_+), \end{aligned} \tag{14}$$

where for every section $i \in I$ and pen $j \in J$:

$$t_i^a = (1-y_i)(t_i+1)+y_i,$$
$$q_{ij}^a = (1-y_i)(q_{ij}-\sum_{k=1}^{q_{ij}} x_{ijk})+y_i q^{\max},$$
$$w_{ij}^a = (1-y_i)w_{ij}+y_i w_0.$$

State s^a is often defined as the *post-decision state*, and s_+ as the *pre-decision state* (Powell 2007). The deterministic transition function $\varphi(s,a)$ only considers the effect of action a while the stochastic transition function $\psi(s^a,\omega)$ takes into account the effect of random information ω. That is, the post-decision state s^a is the state of the system immediately after making the decision while the pre-decision state s refers to the state of the system just before a decision. Figure 1 shows the relation between pre- and post-decision states over three stages.

Now suppose $s \in \mathbb{S}$ is the pre-decision state of the system at the current decision epoch. Moreover, assume that $s^a \in \mathbb{S}$ and $s_-^a \in \mathbb{S}$ are two post-decision states in the current and the previous decision epochs, respectively. According to the transition function $\psi(s_-^a,\omega)$, the pre-decision state s is the result of a stochastic transition from post-decision state s_-^a (see Figure 1) and hence we can conclude that

$$\nu(s_-^a) = \mathbb{E}\left(\nu(s)|s_-^a\right). \tag{15}$$

Furthermore, by reformulating the optimality equations in (2), the value of being in pre-decision state s is

$$\nu(s) = \max_{a\in\mathbb{A}(s)} \left(r(s,a)+\gamma\nu(s^a)\right), \tag{16}$$

where $s^a = \varphi(s,a)$. Now by substituting (16) in (15), the optimality equation for the value function in post-decision state s_-^a becomes

$$\nu(s_-^a) = \mathbb{E}\left(\max_{a\in\mathbb{A}(s)} \left(\, r(s,a)+\gamma\nu(s^a)\right)|s_-^a\right). \tag{17}$$

Note that in (17), the expectation is now outside the max operator and the main benefit of using this optimality equation, compared to (2), is that the optimization problem contained in the expectation is deterministic and hence it can be solved by well-known optimization techniques. However, to calculate $\nu(s_-^a)$, the expectation in equation (17) must be computed. Calculation of this expected value is often computationally intractable and is approximated using simulation techniques. For more details about post-decision states, see Powell (2007, Chapter 4).

3.3 Approximate value iteration algorithm

An *approximate value iteration (AVI)* algorithm is used to estimate the value function (17) using the approximation architecture of the value function defined in (13) and post-decision states. That is,

Algorithm 1 - `AVI` (approximate value iteration algorithm)

```
1  INI: Set $h=1$ and initialize the slope parameters $\hat{b}$ (Pourmoayed et al. 2016);
2  DO
3    Select $s$ randomly; $\hat{b}_{-} := \hat{b}$;
4    FOR $n=1$ TO $N$ DO
5      Solve $a^* = \text{argmax}_{a \in \mathbb{A}(s)} (r(s,a) + \gamma \hat{v}_{\hat{b}}(s^a))$;
6      IF $(n>1)$ THEN $\hat{b}$ :=UpdateSlope($s$, $s^a_{-}$, $h$, $\hat{b}$, $a^*$);
7      $s^{a^*} = \varphi(s,a^*)$;
8      $s := \psi(s^{a^*}, \tilde{\omega})$ (where $\tilde{\omega}$ denotes a sample generated);
9      $s^a_{-} := s^{a^*}$;
10   END FOR
11   $h := h+1$;
12 UNTIL $(n>1)$ AND $(h=H$ OR $||\hat{b}-\hat{b}_{-}|| \leq \lambda)$
```

Algorithm 2 - `UpdateSlope` (update estimates of slope parameters b)

```
1  INPUT: $s=(t,q,w)$, $s^a_{-}=(t^a_{-},q^a_{-},w^a_{-})$, $h$, $\hat{b}$ and $a^*=(x^*,y^*)$.
2  FOR $i \in I$ DO
3    IF ( $t_i \geq t^{\min}$ ) THEN
4      FOR $j \in J$ DO
5        IF ( $q_{ij} < q^{\max}$ ) THEN
6          Set $\tilde{s}=(t,\tilde{q},w)$ where $\tilde{q}_{ij}=q_{ij}+1$;
7          $\nabla\hat{v}_{ij} = \max_{a \in \mathbb{A}(\tilde{s})}(\hat{r}(\tilde{s},a)+\gamma\hat{v}_{\hat{b}}(\tilde{s}^a)) - \max_{a \in \mathbb{A}(s)}(\hat{r}(s,a)+\gamma\hat{v}_{\hat{b}}(s^a))$;
8        ELSE
9          Set $\tilde{s}=(t,\tilde{q},w)$ where $\tilde{q}_{ij}=q_{ij}-1$;
10         $\nabla\hat{v}_{ij} = max_{a \in \mathbb{A}(s)}(\hat{r}(s,a)+\gamma\hat{v}_{\hat{b}}(s^a)) - max_{a \in \mathbb{A}(\tilde{s})}(\hat{r}(\tilde{s},a)+\gamma\hat{v}_{\hat{b}}(\tilde{s}^a))$;
11       END IF
12       $\hat{b}\left((t^a_{-})_i,(w^a_{-})_{ij}\right) := (1-\alpha_h)\hat{b}\left((t^a_{-})_i,(w^a_{-})_{ij}\right) + \alpha_h \nabla\hat{v}_{ij}$;
13     END FOR
14   END IF
15   IF $(y_i{}^*=1)$ THEN
16     $\hat{v} = \frac{\sum_{j \in J} q_{ij}\hat{b}(t_i,w_{ij}) - (\sum_{j \in J} \hat{c}^{\texttt{cull}}(w_{ij})q_{ij} - c^{\texttt{piglet}} q^{\max}|J|)}{\gamma q^{\max}|J|}$;
17     $\hat{b}(1,w_0) := (1-\alpha_h)\hat{b}(1,w_0) + \alpha_h \hat{v}$;
18   END IF
19 END FOR
```

estimates of the slope parameters

$$\hat{b} = \left(\hat{b}\left(t^a_1, w^a_{11}\right), .., \hat{b}\left(t^a_i, w^a_{ij}\right), .., \hat{b}\left(t^a_{|I|}, w^a_{|I||J|}\right)\right),$$

must be found.

The pseudo-code of the algorithm is illustrated in Algorithm 1. Let $\hat{v}_{\hat{b}}(\cdot) : s^a \to \mathbb{R}$ denote the approximate value function given estimates $\hat{b}$. First, the slope parameters are initialized using estimated slope values obtained by considering the model at pen level (Pourmoayed et al. 2016).

Next, the main part of the algorithm is executed using two loops. Each iteration of the outer loop (lines 2-12) simulates a sample path by randomly selecting an initial state s on line 3 and then simulating stages $1, 2, \ldots, N$ using an interior loop.

At each stage, the optimal action is found on line 5 and the slope parameters are updated on line 6. Next, the post-decision state is calculated on line 7 and the pre-decision state for the next stage is found on line 8. Finally, the post-decision state is stored on line 9 because we need it in the sub-algorithm (Algorithm 2) in the next iteration of the loop. The algorithm is stopped after generating at most H sample paths or if the maximum difference between two consecutive estimates of the slopes in the outer loop is less than or equal to a predefined error criterion:

$$||\hat{b}-\hat{b}_{-}|| = \max_{t_i^a \in \{1,..,t^{\max}\}, w_{ij}^a \in \mathbb{W}} \left(|\hat{b}(t_i^a, w_{ij}^a) - \hat{b}_{-}(t_i^a, w_{ij}^a)| \right).$$

More details about the algorithm are described below.

Optimal action Given the pre-decision state s in each iteration, the optimal action $a^* = (x^*, y^*)$ must be found by maximizing $r(s,a) + \gamma \hat{v}_{\hat{b}}(s^a)$ on line 5. The one-stage reward $r(s,a)$ is given in (10) and the value function $\hat{v}_{\hat{b}}(s^a)$ is obtained using the approximation architecture (13) and the deterministic transition of the post-decision state defined in (14). The set of possible actions $\mathbb{A}(s)$ is defined using constraints (3)-(9). Hence the maximization problem becomes:

$$\begin{aligned} \max & \sum_{i\in I}\sum_{j\in J}\sum_{k=1}^{q_{ij}} c_k^{\texttt{cull}}(w_{ij})x_{ijk} - \sum_{i\in I}\sum_{j\in J}\sum_{k=1}^{q_{ij}} c_k^{\texttt{feed}}(w_{ij})(1-x_{ijk}) - c^{\texttt{truck}}z - \\ & \sum_{i\in I} c^{\texttt{piglet}}|J|q^{\max}y_i + \gamma \sum_{i\in I}\sum_{j\in J}\Big((q_{ij} - \sum_{k=1}^{q_{ij}} x_{ijk})\hat{b}\left(t_i^a, w_{ij}^a\right) + y_i q^{\max}\hat{b}(1, w_0)\Big), \qquad (18) \\ & \text{s.t. (3)-(9).} \end{aligned}$$

Sample path Each sample path is simulated over N stages using the interior loop (lines 4-10). Because we cannot simulate an infinite trajectory, parameter N is chosen large enough to give a good estimation of the total expected discounted reward of the model defined in (1). That is, N is chosen such that the discount factor γ^{N-1} becomes sufficiently small and hence further stages have a small impact on the value function (for more details, see (Powell 2007, page 340)). In general, the use of a discount factor close to one means that states in the distant future are assigned the same value of money as are states in the present. Both a VI and an AVI algorithm will therefore converge very slowly.

Updating the slope The sub-algorithm for updating the slope parameters is given in Algorithm 2. The slope parameters are updated using the gradients of the value function with respect to the number of pigs in each pen. That is, given a pen with weight information w_{ij} in week t_i, one pig of average weight is added (line 6) or if there are $q^{\max}$ pigs in the pen, one pig is removed (line 9). This is used to calculate the gradient $\nabla \hat{v}_{ij}$ which, according to (13) and (15), is an estimate of the slope $b((t_{-}^a)_i, (w_{-}^a)_{ij})$ at post-decision state s_{-}^a.

Note that the slope parameters in value function (13) are dependent on the weight information at pen level (not on the weights of the individual pigs). Hence for estimating the gradient, we consider an average weight for the pigs in the pen and therefore the one-stage reward (10) in the maximization problem becomes:

$$\hat{r}(s,a) = \sum_{i\in I}\sum_{j\in J}\sum_{k=1}^{q_{ij}} \hat{c}^{\texttt{cull}}(w_{ij})x_{ijk} - \sum_{i\in I}\sum_{j\in J}\sum_{k=1}^{q_{ij}} \hat{c}^{\texttt{feed}}(w_{ij})(1-x_{ijk}) - c^{\texttt{truck}}z - \sum_{i\in I} c^{\texttt{piglet}}|J|q^{\max}y_i,$$

where $\hat{c}^{\texttt{cull}}(w_{ij})$ and $\hat{c}^{\texttt{feed}}(w_{ij})$ denote the reward and feeding cost of a pig with average weight. This ensures a more stable convergence of the gradients. The slope parameter is updated on line 12 where the step size α_h specifies the weight of the current estimate when updating the slope at iteration h. The step size α_h has an important effect on the convergence rate of the algorithm and our computational testing showed that the *generalized harmonic step size* $\alpha_h = \bar{\alpha}/(\bar{\alpha}+h-1)$ performs well (Powell 2007, page 430). Note that we only update slope parameters if $t_i \geq t^{\min}$ because when $t_i < t^{\min}$, all terms related to section i in the objective function (18) are constant ($x_{ijk} = 0$ for all $j \in J$ and $k \in \{1,2,...,q_{ij}\}$) and hence do not have an impact on the solution.

If termination occurs in section i ($y_i^* = 1$), we update the slope $b(1,w_0)$ (lines 16-17) by considering a contribution to the approximate value function for section i which according to (13) is

$$\hat{v}_{\hat{b}}(s_i) = \sum_{j\in J} q_{ij}\hat{b}(t_i,w_{ij}), \tag{19}$$

where s_i denote the state of section i (state s may be stated as $(s_1,\dots,s_{|I|})$). Based on (16) this contribution is

$$\hat{v}_{\hat{b}}(s_i) = r(s_i,a_i^*) + \gamma\hat{v}_{\hat{b}}(s_i^{a_i^*}), \tag{20}$$

where a_i^* denote the optimal action ($y_i^* = 1$; $x_{ijk}^* = 1, \forall j\in J, k = 1,\dots q_{ij}$) and

$$r(s_i,a_i^*) = \sum_{j\in J} \hat{c}^{\texttt{cull}}(w_{ij})q_{ij} - c^{\texttt{piglet}}q^{\max}|J|,$$

denote the reward related to terminating section i. Hence by combining (19) and (20) we get:

$$\sum_{j\in J} q_{ij}\hat{b}(t_i,w_{ij}) = r(s_i,a_i^*) + \gamma v(s_i^{a^*}) = \sum_{j\in J} \hat{c}^{\texttt{cull}}(w_{ij})q_{ij} - c^{\texttt{piglet}}q^{\max}|J| + \gamma|J|q^{\max}b(1,w_0),$$

and an estimate for $b(1,w_0)$ (see line 16) is

$$\hat{v} = \frac{\sum_{j\in J} q_{ij}\hat{b}(t_i,w_{ij}) - (\sum_{j\in J}\hat{c}^{\texttt{cull}}(w_{ij})q_{ij} - c^{\texttt{piglet}}q^{\max}|J|)}{\gamma|J|q^{\max}}.$$

When the AVI algorithm terminates, the estimated slope parameters can be used to find the best actions in the system. For an arbitrary observed state s, the best action a^* can be found by solving

$$a^* = \arg\max_{a\in\mathbb{A}(s)} (r(s,a) + \gamma\hat{v}_{\hat{b}}(s^a)).$$

Table 1: Parameter values used in the base scenario of the model.

Parameter	Value	Description
$q^{\max}$	18	Maximum number of pigs in a pen.[a]
$t^{\max}$	15	Maximum number of weeks in a growing period (week).[a]
$t^{\min}$	9	First possible week of marketing decisions.[a]
$\lvert I \rvert$	3	Number of sections.[a]
$\lvert J \rvert$	20	Number of pens in a section.[a]
$k^{\texttt{truck}}$	205	Capacity of truck transferring the culled pigs to the abattoir (pig).[b]
$c^{\texttt{truck}}$	400	Fixed cost of the truck (DKK).[b]
$c^{\texttt{piglet}}$	375	Cost of buying a piglet (DKK).[c]
γ	0.95	Discount factor.
w_0	$(26.4, 3, 6.2)$	Initial weight information, $w_0 = (\mu_0, \sigma_0, g_0)$, at the start of a production cycle (kg).[d]
$p^{\texttt{feed}}$	1.8	Unit cost of feed-mix (DKK/FEsv[e]).[f]
$p^{\texttt{pork}}$	10.8	Maximum price of 1 kg carcass (DKK/kg).[f]

[a] Value based on discussions with experts in Danish pig production. [b] Value taken from abattoir (`http://www.danishcrown.dk/ejer/svineleverandoer/danish-crown-afregning/danish-crown-logistik-svin/`). [c] Value taken from Kristensen et al. (2012). [d] Estimated based on time series generated using simulation. [e] FEsv is the energy unit used for feeding pigs in Denmark and is equivalent to 7.72 MJ. [f] Value used in Appendix B for calculating $c_k^{\texttt{cull}}(w_{ij})$ and $c_k^{\texttt{feed}}(w_{ij})$ based on information from abattoir (`http://www.danishcrown.dk/Ejer/Noteringer/Aktuel-svinenotering.aspx`).

where $s^a = \varphi(s, a)$. That is, in contrast to the usual solution procedures for MDPs, we do not generate a policy containing optimal actions for every possible state. Instead, we solve the above maximization problem to find the best action for the observed states in the system.

4 Computational results

In order to show the functionality of the proposed model, we conduct four numerical experiments. In Experiment 1, we apply the AVI algorithm at pen level to show the accuracy of the algorithm. Because the size of the model is tractable at pen level, we first solve the model using non-approximate value iteration and compare the results with the AVI algorithm. In Experiment 2, we apply the AVI algorithm at herd level to investigate how the best marketing decisions can be found under different average growth rates in the sections. Next, in Experiment 3, the marketing policy obtained by the AVI algorithm is compared with other marketing policies that may be applied to the herd. Finally, in Experiment 4, we carry out a small sensitivity analysis for some parameters in the model.

4.1 Parameters

The values of the parameters used in the model are given in Table 1. These values have been obtained using information about finisher pig production units (Danish conditions) and the related literature. Moreover, the length of the cleaning period (Assumption 7 on Page 6) and the lead time for ordering a truck (Assumption 3 on Page 6) have been set to 4 and 3, respectively.

Table 2: Range of the centre points for continuous state variables μ_{ij}, σ_{ij}, and g_{ij} (kg). The continuous state variables μ_{ij}, σ_{ij}, and g_{ij} are discretised into 21, 9, and 6 intervals, respectively.

Week (t)	1	2	3	4	5	6	7	8
μ	6.4-46.4	13.4-53.4	20.4-60.4	27.4-67.4	34.4-74.4	41.4-81.4	48.4-87.4	55.4-95.4
σ	2-10	2.5-10.5	3-11	3.5-11.5	4-12	4.5-12.5	5-13	5.5-13.5
g	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2
Week (t)	9	10	11	12	13	14	15	
μ	62.4-102.4	69.4-109.4	76.4-115.4	83.4-123.4	90.4-130.4	97.4-137.4	104.4-144.4	
σ	6-14	6.5-14.5	7-15	7.5-15.5	8-16	8.5-16.9	9-17	
g	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2	4.2-8.2	

In the model, we also need the functions $c_k^{\texttt{cull}}(w_{ij})$ and $c_k^{\texttt{feed}}(w_{ij})$ used in the reward function (10). These functions depend on the carcass weight, leanness, feeding cost, and the piecewise linear settlement pork price of each pig in the pen. All functions are described in detail in Pourmoayed et al. (2016) and hence we only provide a short overview in Appendix B. Moreover, because the framework of the model is based on a discrete-state MDP, the continuous state variables μ_{ij}, σ_{ij}, and g_{ij} related to weight information w_{ij} (see Section 2.1) must be discretised into a set of intervals. We use the same method as in Pourmoayed et al. (2016). Thereby, in every week of the growing period the state variables μ_{ij}, σ_{ij}, and g_{ij} are divided into 21, 9, and 6 intervals, respectively. That is, interval lengths of 2 kilograms, 1 kilograms and 0.8 kilograms, respectively. Hence, the size of the set $\mathbb{W}$ representing possible weight information is $|\mathbb{W}| = 21 \cdot 9 \cdot 6 = 1,134$. An overview of the discretisation is given in Table 2 where the ranges of the centre points of each interval are shown. These ranges represent realistic values of weight and growth information in each week of the growing period. For further information, see Pourmoayed et al. (2016).

The AVI algorithm was coded in C++ (MS VS2010 compiler) using CPLEX 12.6.2 (C++ API by Concert Technology) as the optimization solver to solve the maximization problem of the algorithm. The source code is available online (Pourmoayed 2016). In the algorithm, parameters N, H, λ, $\bar{\alpha}$ are set to 120, 300, 5, and 100, respectively. The best setting was obtained by testing the algorithm with different values of these parameters. Note that with these parameter values, at most 36,000 decision epochs are simulated in the algorithm and in each simulation the slope parameters are updated according to the weight information in 60 pens, i.e., the maximization problem of the algorithm is solved up to 2,160,000 times in total.

The sample paths used in the computational experiments are randomly generated from the stochastic process describing the evolution of weight information $w_{ij} = (\mu_{ij}, \sigma_{ij}, g_{ij})$ in the pens (for more details see Appendix A). This process depends on the conditional probability distributions of the *state space models* formulated in Pourmoayed et al. (2016).

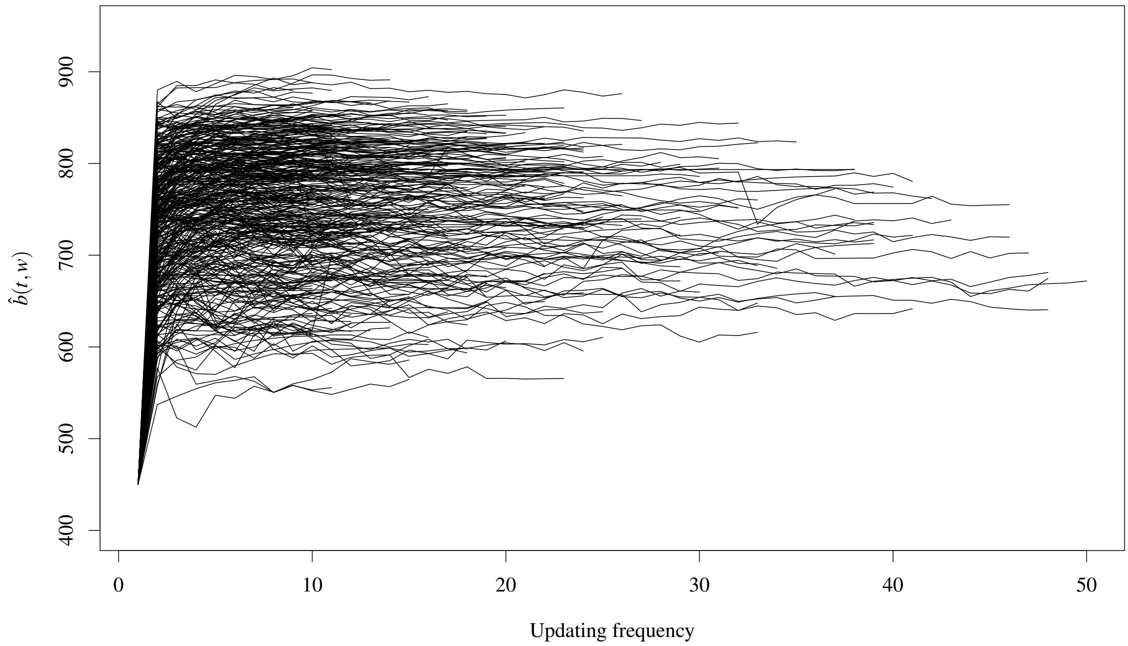

Figure 2: Convergence of the slope parameters in the AVI algorithm. The horizontal axis shows the number of updates of an observed slope, and the vertical axis indicates the estimated value of the observed slope. The lines represent different combinations of week and weight information.

4.2 Experiment 1: Accuracy of the AVI algorithm

In this section, we investigate the accuracy of the AVI algorithm and compare it with the value iteration (VI) algorithm. To have a tractable model, we consider a single section with one pen and use the VI algorithm to solve the MDP model under the total discounted reward criterion described in Section 2.

In order to compare the AVI and VI algorithms, we compare the slopes resulting from the VI algorithm with the estimated slopes obtained by the AVI algorithm. The AVI algorithm was initialized using a value of 450 for the slope parameter for all combinations of t and w and terminated after approximately 45 minutes. Figure 2 illustrates the convergence of the slope parameters obtained by the AVI algorithm. Because the algorithm is based on sample paths, some of the states are observed more often than others and hence the related slope parameters converge faster.

Figure 3 compares the slope values obtained by the AVI algorithm against the slope values obtained by the VI algorithm. Each point corresponds to the slope values for a specific pair (t,w) such that the x-coordinate value of a point is the estimated value of slope $b(t,w)$ obtained by the AVI algorithm and the y-coordinate is the slope value estimated by the VI algorithm. The dashed line is the 45-degree line where the slope values of x- and y-coordinate are equal. Figure 3 shows that the data points are relatively close to the dashed line indicating that there is a little difference between the

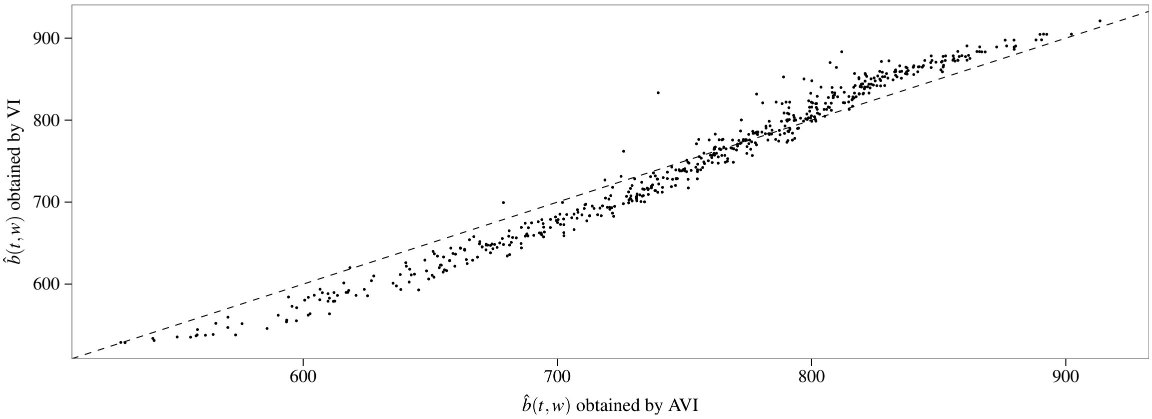

Figure 3: Comparison between the estimated slope values obtained by the AVI and VI algorithms. Each point corresponds to a specific pair (t,w). The x-coordinate value of each point is the estimated value of slope $b(t,w)$ obtained by the AVI algorithm, and the y-coordinate value is the slope value estimated by the VI algorithm. The dashed line is a 45-degree line considered as a marker in the plot when the values of the x and y coordinates are equal.

slope values calculated by the AVI and VI algorithms. However, underestimation and overestimation of a slope may happen. A closer analysis of the data showed that this happens mainly in the first and last weeks of the growing period when the pen is almost full and empty, respectively. The *root mean square error* (RSME) between the slopes obtained by the AVI and VI algorithms is 22.7, which is a relatively small error with respect to the range of the estimated values (between 500 and 900). As a result, the AVI algorithm gives a reasonable estimation for the slope parameters used in the approximate value function. Note that if the values of parameters N and H are increased in the AVI algorithm, then the estimation error will be lower. However, the parameters of the algorithm have been adjusted in a way such that we have a good balance between the computational efficiency of the algorithm and the quality of the approximated value function. Moreover, when the algorithm is applied at herd level, the approximated value function will be significantly better because the slope parameters are estimated using sample paths with simulated weight information of 60 pens instead of one pen, i.e., each loop of the AVI algorithm for updating the slope parameters will contain 60 interior loops, one for each pen (see Algorithm 2).

4.3 Experiment 2: Marketing decisions at herd level

In this section, we use the AVI algorithm to find the best marketing decisions in a herd with 3 sections. In order to show how the marketing decisions change according to the conditions of the herd, we assume that there are different environmental effects (e.g., temperature, housing conditions, and humidity) for each section resulting in different average growth rates of the pigs in the three sections. The average growth rate of pigs in Section II is considered normal (6 kilograms per week) whereas in Sections I and III, the pigs grow 10 percent slower and faster than in Section II, respectively. The

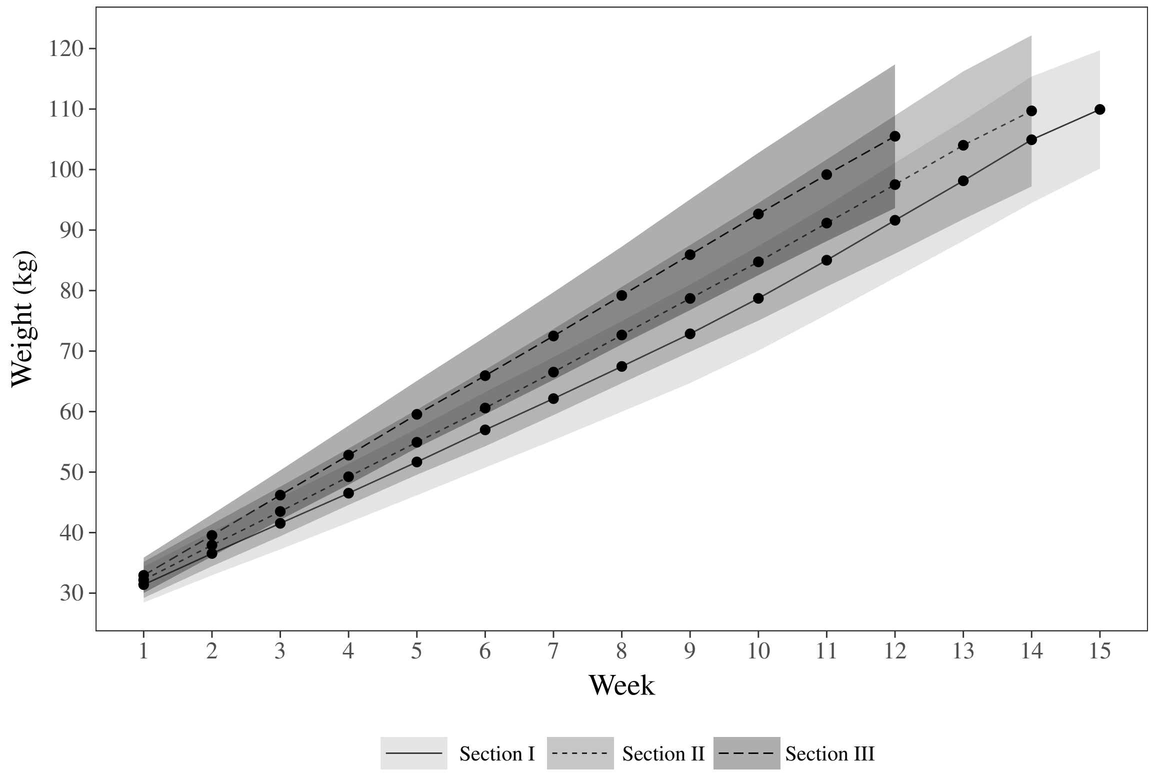

Figure 4: The range of the simulated weight data in the three sections during $t^{\max} = 15$ weeks of growing period in the production cycles. The lines are the average weight of the pigs during the growing period. The ribbon for each section shows the variability of data using one standard deviation of uncertainty.

AVI algorithm was initialized using the slopes obtained from the VI algorithm in Experiment 1 and the parameters in Section 4.1. The algorithm terminated after approximately 21 hours.

Figure 4 shows the range of the simulated weight data during $t^{\max} = 15$ weeks of the growing period in the three sections. The lines show the average weight of the pigs during the growing period in the three sections. The variability of weight data in each section is represented using ribbon surfaces showing one standard deviation of uncertainty. Figure 4 illustrates, in each week (e.g., week 10), the weight values of the pigs in Section III are higher than those in Sections I and II, and the pigs in Section II grow faster compared to Section I. Moreover, when the pigs become older, the weights vary more widely between the pigs of the same age across all sections (see the area of the ribbons from Week 1 to Week 15).

Figure 5 illustrates the best marketing decisions during 49 weeks of production in the herd (from Week 46 to Week 94). We have randomly selected this period to have a plot of marketing decisions in each of the sections which are in different stages (weeks) of the production cycle. That is, weeks 1 to 45 can be considered as the warm-up period of the simulation. In this figure, the bars show the number of remaining pigs before a decision is made, the numbers below the bars denote the number

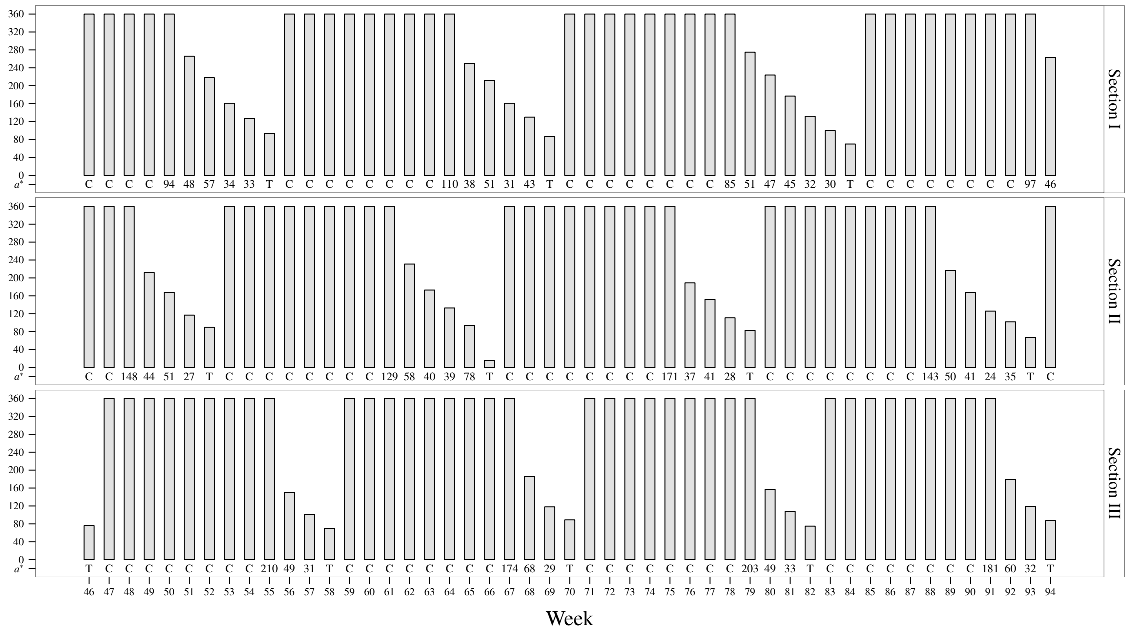

Figure 5: Marketing decisions in a herd with three sections during 49 weeks of production (Week 46 to Week 94). Bars show the number of remaining pigs in the sections before a decision is made. Numbers below the bars denote the number of the heaviest pigs culled from the section, the letter "T" indicates termination of a production cycle in the section, and the letter "C" corresponds to continuing the production process without any marketing.

of the heaviest pigs culled from the section, the letter "T" indicates termination of the production cycle in the section, and the letter "C" corresponds to continuing the production process without any marketing. Note that when termination occurs in a section, a new production cycle is started in this section. Moreover, the time period between two successive terminations shows the length of a production cycle in the section.

As seen in Figure 5, the length of the production cycles in Section III is shorter than in other sections because the average weight of the pigs in this section is higher than in Sections I and II, and hence the pigs obtain their optimal slaughter weight earlier. It is therefore beneficial to terminate the section after a few weeks of individual marketing resulting in a higher number of production cycles in Section III compared to the other sections. In Sections I and II, the pigs grow more slowly than in Section III, and hence it is better to keep them for a longer period compared to Section III, i.e., the lengths of the production cycles in Sections I and II are longer than in Section III. However, after starting the marketing decisions in Sections I and II, we observe that in each production cycle, the fraction of pigs remaining in Section I is higher than in Section II (compare the height of the bars for Sections I and II in Figure 5). This happens because the average growth of pigs in Section II is higher than in Section I and hence more pigs are culled from Section II compared to Section I resulting in a lower number of pigs in Section II at termination.

Figure 6 illustrates the total number of culled pigs in the herd from Week 46 to Week 94. The

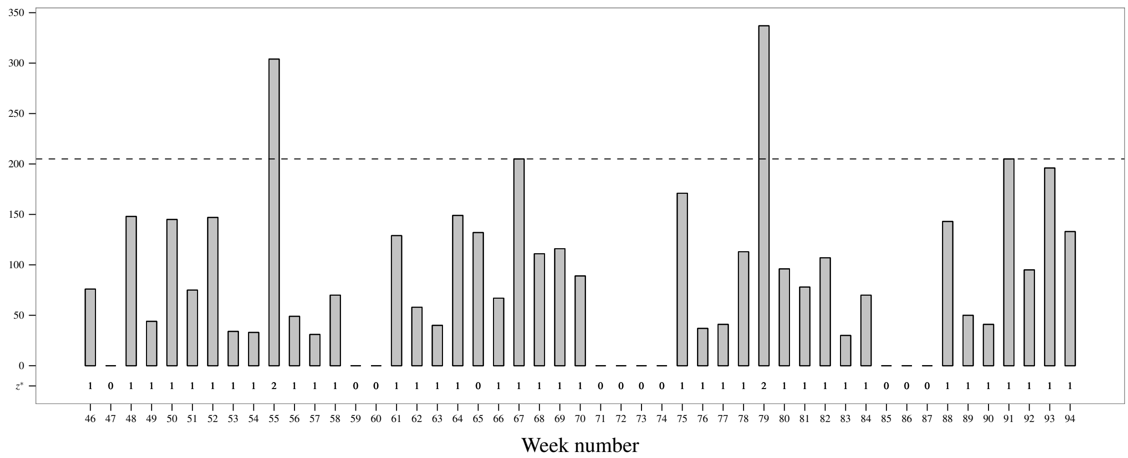

Figure 6: The total number of culled pigs from the herd in Week 46 to Week 94. The numbers below the bars show the number of trucks needed to transfer the culled pigs to the abattoir. The horizontal dashed line shows the full capacity of a truck.

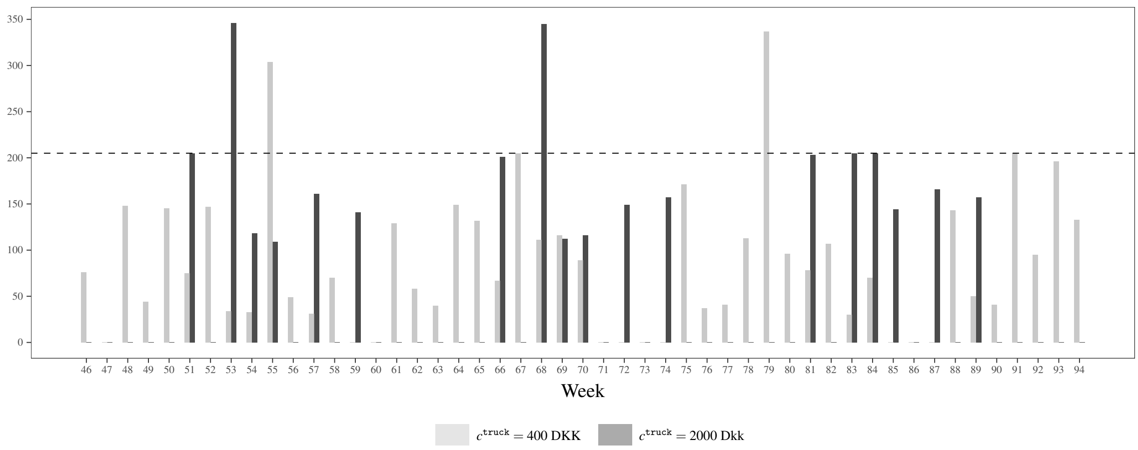

Figure 7: The effect of fixed transportation cost on the marketing policy of the herd. Black and grey bars show the number of culled pigs when the transportation costs are 2000 DKK and 400 DKK, respectively. The horizontal dashed line shows the full capacity of a truck (205 pigs).

numbers below the bars show the number of trucks needed to transport the culled pigs to the abattoir and the horizontal dashed line shows the capacity ($k^{\mathtt{truck}} = 205$) of a single truck. As can be seen, the full capacity of the trucks is not used in most cases and only the pigs with appropriate live weight are sent to the abattoir. This happens because the fixed cost of a truck (400 DKK) is lower than the profit obtained by culling a specific number of pigs that are ready for slaughter. For instance, when 40 pigs are ready for slaughter, it is beneficial to call a truck and transport them to the abattoir. Moreover, in some cases a compulsory termination at the latest week of the growing period (week $t^{\max} = 15$) leads to a small number of pigs being culled and sent to the abattoir by one truck (e.g., Week 84 in Figure 5).

Transportation costs can affect the marketing policy of the farm. Figure 7 shows the change in the number of marketed pigs in the herd when the fixed transportation cost ($c^{\texttt{truck}}$) is changed from 400 to 2,000 DKK. We note that, a high transportation cost leads to better utilization of the trucks and fewer deliveries to the abattoir. More precisely, with a transportation cost of 400, the number of deliveries to the abattoir is 39 while with a transportation cost of 2,000, this number is decreased to 18 and the average cycle length has increased 1.7 weeks. A high transportation cost may result in culling some pigs that have not reached their optimal slaughter weight because the lost revenue from slaughtering at a sub-optimal weight is outweighed by the fixed transportation cost. Note that the variable costs of transportation per pig (e.g., cost of loading a pig into the truck) are not considered in the model. Because the total number of pigs culled per production cycle is fixed, the total variable cost per cycle will also be fixed and hence the variable transportation cost will not have a noticeable impact on the marketing policy.

4.4 Experiment 3: AVI policy compared to other marketing policies

In order to evaluate the performance of the marketing policy obtained by the ADP, we compare it with other well-known marketing polices that are often applied at herd level. We compare the marketing policy calculated with the AVI algorithm with the following policies:

Myopic (M) This policy does not consider the impact of the present decisions on the future conditions of the system, i.e., only the immediate reward of the decisions is taken into account in the model. To find the best marketing decisions under this policy, we solve

$$max_{a\in\mathbb{A}(s)}(r(s,a))$$

at each stage without considering the future reward. This policy was chosen for comparison in order to stress the importance of using dynamic programming that considers the future value of decisions.

Full truck capacity (FTC) Here the farmer prefers to use the maximum capacity of the trucks when a delivery to the abattoir takes place. In order to evaluate this policy, we change the sign of the transportation constraint (6) to an equality constraint if the total number of pigs above the capacity. Moreover, if the total number of pigs is below the capacity, a compulsory termination at herd level occurs. This policy may be followed by farmers who want to use the maximum capacity of the truck when sending the pigs to the abattoir.

Pen-level (PL) This policy use the marketing decisions found at pen level in Experiment 1. That is, marketing decisions are made without taking into account the cross-level transportation costs. This policy was chosen to illustrate the effect of ignoring herd level considerations.

All-In All-Out (AIAO) This policy does not consider individual marketing decisions in the pens. Instead, when the length of the production cycle in a section equals a specified value, all pigs are marketed from the section in one delivery. We evaluate this policy under the assumption that a section is terminated in week t ($t = 9, \ldots, 15$). This policy is commonly used in the industry as the main delivery strategy to the abattoir.

No learning (NL) Here no learning is applied to the slope parameters, i.e., the slope parameters remain at the values initialized from the pen-level model.

Table 3: Performance of different marketing policies. For each policy, 100 sample paths are applied and results are reported as 95% confidence intervals $m \pm 1.96s/\sqrt{100}$, where m and s are the mean and standard deviation of the values.

Policy	AVI	M	FTC	PL	NL	AIAO						
						9	10	11	12	13	14	15
Discounted reward per week (DKK/herd)	2072 ± 45.5	1078 ± 11.1	1830 ± 103	1745 ± 13	1025 ± 23.2	470 ± 11.9	1398 ± 10.3	1831 ± 7.6	1881 ± 6.7	1679 ± 4.6	1365 ± 4	1014 ± 3.9
Discounted reward reduction (%)	0	48.0	11.7	15.8	50.6	77.3	32.5	11.6	9.2	19.0	34.1	51.1
Length of production cycle (weeks)	11.8 ± 0.01	10 ± 0	11.6 ± 0.06	10.7 ± 0.01	15 ± 0	9 ± 0	10 ± 0	11 ± 0	12 ± 0	13 ± 0	14 ± 0	15 ± 0
Number of production cycles[a]	30.5 ± 0.14	36 ± 0	31.3 ± 0.28	34.4 ± 0.16	24 ± 0	39 ± 0	36 ± 0	33 ± 0	30 ± 0	27 ± 0	27 ± 0	24 ± 0
% of truckload capacity utilized	76 ± 1	75 ± 0	88 ± 0	69 ± 0	88 ± 0	87 ± 0	87 ± 0	87 ± 0	87 ± 0	87 ± 0	87 ± 0	87 ± 0
Number of trucks sent to abattoir[b]	73.7 ± 1	84 ± 0	63.6 ± 0.5	89.3 ± 0.9	48 ± 0	84 ± 0	72 ± 0	66 ± 0	60 ± 0	54 ± 0	54 ± 0	48 ± 0

[a] Total number of production cycles in all three sections over 120 weeks. [b] Total number of trucks sent to the abattoir over 120 weeks.

To compare marketing policies, we generated 100 sample paths over a period of 120 weeks (approximately 2.5 years) in a herd where all sections follow an average growth rate of 6 kilograms per week which are used to evaluate all the policies. Moreover, the following statistics are calculated for each policy π: discounted reward per week defined as $\sum_{n=1}^{N} \gamma^{n-1} r_n(s, \pi(s))/N$, length of a production cycle, number of production cycles, percentage of truckload capacity utilized, and number of trucks sent to the abattoir.

Table 3 shows the results for the AVI, M, FTC, PL, NL and AIAO policies. Results are reported as 95% confidence intervals. The policy found by the AVI algorithm outperforms other policies in terms of average discounted reward per week, and the AIAO policy with termination in week 9 results in the lowest reward.

For $c^{truck} = 400$, the difference between the ADP and FTC policies is noticeable. Using the maximum capacity of trucks in the FTC policy results in marketing some pigs that are not ready for slaughter yet and hence the farmer loses the possible profit that could have been earned by keeping these pigs for a longer time in the herd. Note that in this policy, as expected, the utilization of truckload capacity is better than in other policies. Using a myopic policy results in significant losses because we do not take into account the effect of current decisions on future rewards.

When considering the AIAO policies, it seems that policies with a length of 11 or 12 weeks perform well. Note that the length of the production cycle in these policies is close to the average length of the production cycle under the AVI policy. That is, the difference in discounted average reward is mainly due to individual marketing that is possible under the AVI policy.

Because the PL policy does not take into account the cross-level transportation cost, the utilization of truckload capacity is low compared to other policies. This highlights the importance of considering marketing decisions at herd level. Using optimal decisions at pen level may result in higher transportation costs.

In the AIAO policies, as expected, when the length of the growing period is increasing (from 9 to 15 weeks), the number of production cycles decreases during the 120 weeks of production and hence we need fewer trucks to transfer the culled pigs to the abattoir. Furthermore, in these policies the utilization of truckload capacity is high and close to the FTC policy, because all the pigs are sent to the abattoir in specific weeks and hence the capacity of most trucks is utilized.

Using the initial slope parameters from the pen-level model when finding optimal actions by solving model (18) finds a policy very close to the AIAO policy (15 weeks). This demonstrates the effectiveness of the AVI's learning.

Finally, note that for longer production cycles (e.g., 14 or 15 weeks), the average discounted reward per week is reduced noticeably. This happens because the feeding cost of the pigs increases considerably during the last weeks of the growing period, and the revenue obtained by selling the heavy pigs cannot compensate for the increased feeding cost. Hence the average reward is reduced which demonstrates the impact of feeding costs.

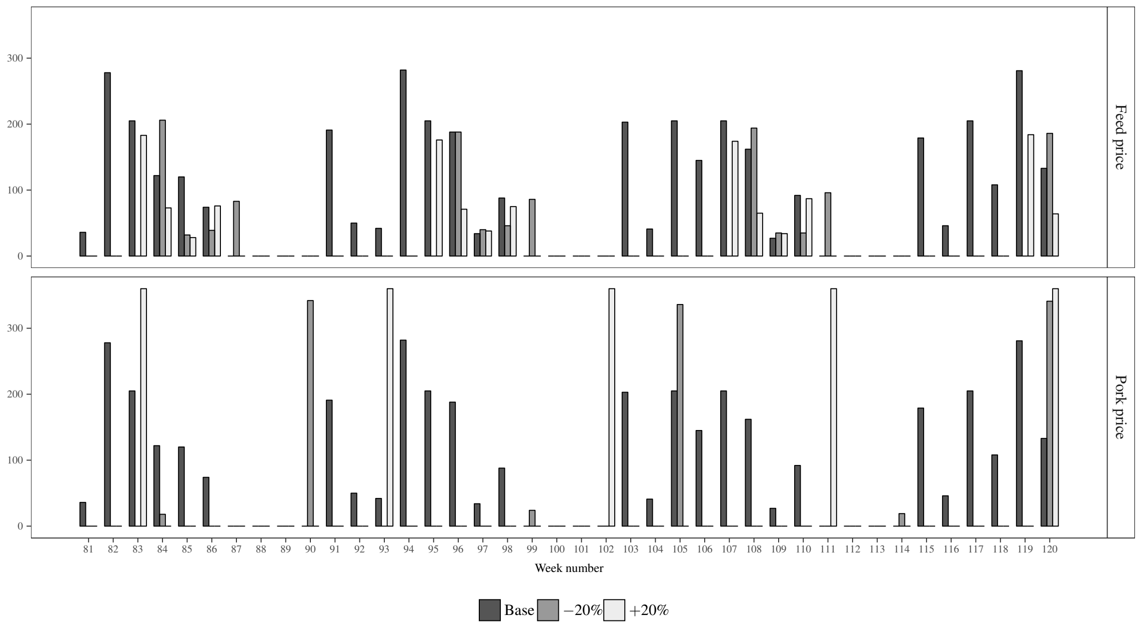

Figure 8: The effect of marginal changes in the unit feed or pork price on the total number of culled pigs in Section II of the herd.

Table 4: Percentage effect of a 20% marginal increase/decrease in the unit feed-mix or pork price compared to the base scenario (the parameters in Table 1).

	Feed price		Pork price	
	-20%	+20%	-20%	+20%
Discounted reward per week (DKK/herd)	52.8	-37.1	-165.5	91.9
Length of production cycle (weeks)	1.1	-1.0	26.7	-23.9
Number of production cycles[a]	-1.3	2.3	-21.2	28.1
% of truckload capacity utilized	-0.5	-2.2	4.2	16.1
Number of trucks sent to abattoir[b]	-2.2	4.3	-27.7	13.6

[a] Total number of production cycles in all three sections over 120 weeks. [b] Total number of trucks sent to the abattoir over 120 weeks.

4.5 Experiment 4: Sensitivity analysis

In order to see the effect of changes in feed and pork prices we conduct a small sensitivity analysis. We compare the AVI policy calculated in Experiment 2 against a 20% marginal increase/decrease in the unit feed-mix or pork price in a herd with three sections and different average growth rates. Similar to Experiment 3, we calculate a set of key performance indicators using 100 sample paths over a period of 120 weeks in the herd.

The percentage effect of a 20% marginal increase/decrease in the unit feed-mix price or pork price compared to the base scenario is shown in Table 4. Moreover, the effect of marginal changes on the total number of culled pigs in Section II of the herd is illustrated in Figure 8.

As expected, a decrease (increase) in the feed-mix cost and pork price gives a higher (lower) average discounted reward. Moreover, as seen in the table, the pork price is much more sensitive compared to the feed-mix unit price. That is, the objective of the model is more sensitive to changes in the pork price.

A 20% decrease in the pork price results in a negative reward (a decrease below 100%). The farmer may of course close the production in such a case. However, in practice often this is not possible. In this case it is best to keep the pigs in the pens as long as possible resulting in a cycle length of 15 weeks as shown in Figure 8. However, a few pigs with a very high weight (about 90-100 kilogram) are culled at week 9 in a production cycle because otherwise the feeding cost would be very high. On the contrary, if the pork price increases by 20%, the price is so favourable that most pigs are culled as soon as possible (-23.9% $\approx$ 9 weeks). Figure 8 demonstrates this accelerated marketing where a section of 360 pigs are culled with a gap of 9 weeks in-between. Both cases result in culling pigs at the lower or upper bounds of the length of a production cycle and hence almost no individual marketing is done. That is, batches sent to the abattoir are larger (4.2% and 16.1%).

A decrease in the feed price results in a longer cycle length because it is more optimal to feed the pigs for a longer period. As a result larger batches are sent to the abattoir. An increase in the feed price have a negative impact on reward and the cycle length is reduced with 1%. Individual marketing increase, resulting in lower truck utilization (-2.2%).

5 Conclusions

In this paper, we consider sequential marketing decisions of finishing pigs at herd level and take into account cross-level constraints (termination at section level and transportation at herd level).

The problem is modelled using a discounted infinite-horizon MDP where the state of the system is based on the weight information of each pen in the herd and is updated using state space models and a Bayesian approach. Due to the curse of dimensionality, we apply an ADP approach and solve the model by using an AVI algorithm based on post-decision states to find the best marketing policy. The value function of the model is approximated by a parametric function and the AVI algorithm estimates the slope parameters.

A set of numerical experiments was conducted. First, by comparing the AVI algorithm with the exact solution at pen level, it seems that good estimates of slope parameters can be found by the AVI algorithm. Second, the AVI algorithm was used to find the best marketing decisions in a herd under different conditions of average growth in the sections. It was illustrated how the marketing policy adapts under different fixed transportation costs. Increasing transportation cost results in higher truck utilization and longer production cycles. Next, the best policy found by the AVI algorithm was compared against several well-known marketing policies. This policy outperformed other marketing polices in terms of average discounted reward per week. Having individual marketing, i.e., allowing multiple shipping dates within a section will increase profit. This result is in agreement with the results in Boys et al. (2007) and Ohlmann and Jones (2011). Optimizing the current production cycle without considering the effect on the future production cycles (a myopic policy) or trying to utilizing the maximum capacity of the truck (FTC policy) is not recommended which also was confirmed in Ohlmann and Jones (2011). Moreover, ignoring herd constraints and culling pigs based on an optimal pen-level policy will reduce profit. This highlights the importance of considering the whole herd when optimizing marketing. Finally, a sensitivity analysis showed the importance of feed and pork prices on the reward of the model.

The model presented in this paper has an operational focus and optimizes decisions sequentially given the current state of the system. The uncertainty is represented using state space models. That is, a Bayesian approach is used to update the system such that it contains information based on the previous data. The state space models are embedded into the MDP such that the state of each pen is found using the current data stream for each pen in the herd. Hence the sectioning of the production facility is considered because the model optimizes sequential decisions based on information about each pen in the herd.

Acknowledgments

The authors would like to thank Professors Anders Ringgaard Kristensen and Christian Larsen for useful comments on an earlier draft of this paper. Moreover, we thank an anonymous reviewer for careful reading the manuscript and his many insightful comments and suggestions. This article has been written with support from The Danish Council for Strategic Research (The PigIT project, Grant number 11-116191).

References

S. Andersen, B. Pedersen, and M. Ogannisian. Slagtesvindets sammensætning. meddelelse 429. Technical report, Landsudvalget for Svin og Danske Slagterier, 1999. URL `http://vsp.lf.dk/Publikationer/Kilder/lu_medd/medd/429.aspx`.

Y. Ben-Ari and S. Gal. Optimal replacement policy for multicomponent systems: An application to a dairy herd. *European Journal of Operational Research*, 23(2):213–221, 1986. doi:10.1016/0377-2217(86)90240-7 .

K.A. Boys, N. Li, P.V. Preckel, A.P. Schinckel, and K.A. Foster. Economic replacement of a heterogeneous herd. *American Journal of Agricultural Economics*, 89(1):24–35, 2007. doi:10.1111/j.1467-8276.2007.00960.x .

D.P. de Farias and B. Van Roy. The linear programming approach to approximate dynamic programming. *Operations Research*, 51(6):850–865, 2003. doi:10.1287/opre.51.6.850.24925 .

S. Gal. The parameter iteration method in dynamic programming. *Management Science*, 35(6):675–684, 1989. doi:10.1287/mnsc.35.6.675 .

J.J. Glen. A dynamic programming model for pig production. *Journal of the Operational Research Society*, 34: 511–519, 1983. doi:10.1057/jors.1983.118 .

E. Jørgensen. The influence of weighing precision on delivery decisions in slaughter pig production. *Acta*

Agriculturae Scandinavica, Section A - Animal Science, 43(3):181–189, August 1993. doi:10.1080/09064709309410163 .

E. Jørgensen. Foderforbrug pr kg tilvækst hos slagtesvin. fordeling mellem forbrug til vedligehold og til produktion i besætninger under den rullende afprøvning. Technical report, Danish Institute of Agricultural Sciences, Biometry Research Unit, 2003.

A.R. Kristensen. Optimal replacement in the dairy herd: A multi-component system. *Agricultural Systems*, 39 (1):1–24, 1992. doi:10.1016/0308-521X(92)90002-6 .

A.R. Kristensen, L. Nielsen, and M.S. Nielsen. Optimal slaughter pig marketing with emphasis on information from on-line live weight assessment. *Livestock Science*, 145(1-3):95–108, May 2012. doi:10.1016/j.livsci.2012.01.003 .

H. Kure. *Marketing Management Support in Slaughter Pig Production*. PhD thesis, The Royal Veterinary and Agricultural University, 1997. URL `http://www.prodstyr.ihh.kvl.dk/pub/phd/kure_thesis.pdf`.

J.K. Niemi. *A dynamic programming model for optimising feeding and slaughter decisions regarding fattening pigs | NIEMI | Agricultural and Food Science*. PhD thesis, MTT Agrifood research, 2006. URL `http://ojs.tsv.fi/index.php/AFS/article/view/5855`.

J.W. Ohlmann and P.C. Jones. An integer programming model for optimal pork marketing. *Annals of Operations Research*, 190(1):271–287, November 2011. doi:10.1007/s10479-008-0466-3 .

L.M. Plà-Aragonés, S. Rodriguez-Sanchez, and V. Rebillas-Loredo. A mixed integer linear programming model for optimal delivery of fattened pigs to the abattoir. *Journal of Applied Operational Research*, 5:164–175, 2013. URL `http://orlabanalytics.ca/jaor/archive/v5n4.htm`.

R. Pourmoayed. Github repository: An approximate dynamic programming approach for sequential pig marketing decisions at herd level, 2016. URL `https://github.com/pourmoayed/adpProject`.

R. Pourmoayed and L.R. Nielsen. An overview over pig production of fattening pigs with a focus on possible decisions in the production chain. Technical Report PigIT Report No. 4, Aarhus University, 2014. URL `http://pigit.ku.dk/publications/PigIT-Report4.pdf`.

R. Pourmoayed, L.R. Nielsen, and A.R. Kristensen. A hierarchical Markov decision process modeling feeding and marketing decisions of growing pigs. *European Journal of Operational Research*, 250(3):925–938, 2016. doi:10.1016/j.ejor.2015.09.038 .

W.B. Powell. *Approximate Dynamic Programming: Solving the curses of dimensionality*, volume 703. Wiley-Interscience, 2007. ISBN 978-0-470-60445-8.

W.B. Powell. Merging AI and OR to solve high-dimensional stochastic optimization problems using approximate dynamic programming. *INFORMS Journal on Computing*, 22(1):2–17, 2010. doi:10.1287/ijoc.1090.0349 .

W.B. Powell and B. Van Roy. *Approximate Dynamic Programming for High-Dimensional Resource Allocation Problems*, chapter 10, pages 261–283. Wiley-Blackwell, 2012. ISBN 9780470544785. doi:10.1002/9780470544785.ch10 .

M.L. Puterman. *Markov Decision Processes: Discrete stochastic dynamic programming*. John Wiley & Sons, 2005. ISBN 978-0-471-72782-8.

N. Toft, A.R. Kristensen, and E. Jørgensen. A framework for decision support related to infectious diseases in slaughter pig fattening units. *Agricultural Systems*, 85(2):120–137, August 2005. doi:10.1016/j.agsy.2004.07.017 .

H. Topaloglu and W.B. Powell. Dynamic-programming approximations for stochastic time-staged integer multicommodity-flow problems. *INFORMS Journal on Computing*, 18(1):31–42, 2006. doi:10.1287/ijoc.1040.0079 .

D. Wackerly, W. Mendenhall, and R. Scheaffer. *Mathematical Statistics with Applications*. Thomson Learning, 2008. ISBN 0495385069.

M. West and J. Harrison. *Bayesian Forecasting and Dynamic Models (Springer Series in Statistics)*. Springer-Verlag, February 1997. ISBN 0387947256.

A Modelling the pen-level pig weight information

Consider a pen with weight information $X = (X_1, X_2) = ((\mu, g), \sigma)$ where $X_1 = (\mu, g)$ and $X_2 = \sigma$ denote the average weight and weekly growth, and the standard deviation of the weights, respectively. Moreover, let $\hat{X}$ denote a realization (sample) of this random variable.

A discrete-time stochastic process $\{X_t, t = 1, 2, ..., t^{\max}\}$ modelling weight information in the pen is defined as a collection of random variables X indexed by week number t. The random variable X_t depends on the earlier values observed in the process, $\hat{X}_{t-1}, \hat{X}_{t-2}, .., \hat{X}_1$. Therefore, in order to generate samples, we need to find the probability distribution of the conditional random variables $X_t|\hat{X}_{t-1}, \hat{X}_{t-2}, .., \hat{X}_1$. In order to find this conditional distribution, we use two *state space models* (SSMs) formulated in Pourmoayed et al. (2016) for modelling the dynamics of the state variables (μ_t, g_t) and σ_t in a finisher pen.

Given d weight measurements at time t measured by e.g., an online sensor and let $\bar{w}_t$ denote the average of these measurements. Moreover assume $\bar{z}_t$ is the average feed intake at time t that can be measured by a feeding control system. The following Gaussian SSM is used to model the mean weight and growth (X_1) (West and Harrison 1997):

$$\text{system equation } (\theta_t = G\theta_{t-1} + \omega_t): \quad \begin{pmatrix} \mu_t \\ g_t \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ g_{t-1} \end{pmatrix} + \begin{pmatrix} \omega_{1t} \\ \omega_{2t} \end{pmatrix},$$

$$\text{observation equation } (y_t = F_t'\theta_t + \nu_t): \quad \begin{pmatrix} \bar{w}_t \\ \bar{z}_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ k_{1t} & k_2 \end{pmatrix} \begin{pmatrix} \mu_t \\ g_t \end{pmatrix} + \begin{pmatrix} \nu_{1t} \\ \nu_{2t} \end{pmatrix}.$$

The system equation of this SSM describes the dynamics of latent variables μ_t and g_t corresponding to average weight and growth of pigs at time t. The observation equation models the relation between latent variables and observable variables $\bar{w}_t$ and $\bar{z}_t$. Here ω_t and ν_t are two random terms in system and observation equations following normal distributions with mean zero and unknown variances V and W, respectively. Variances V and W are estimated using maximum likelihood estimation (see Pourmoayed et al. (2016)). Parameters k_{1t} and k_2 are two constants used to estimate the average feed intake ($\bar{z}_t$) given weight and growth information (μ_t, g_t) (Jørgensen 1993).

In order to model the inhomogeneity among the pigs (X_2), a non-Gaussian SSM is used to model the evolution of the variance of the weights. Assume d weight measurements that are normal distributed with true variance σ_t^2 and let s_t^2 denote the sample variance of these measurements. Since $(d-1)s_t^2/(\sigma_t)^2$ follows a chi-square distribution with $d-1$ degrees of freedom (Wackerly, Mendenhall, and Scheaffer 2008, p357), an SSM can be defined with latent variable σ_t^2 and observable variable $y_t = s_t^2$ that follows a gamma distribution with shape and scale parameters $\frac{d-1}{2}$ and $\frac{2(\sigma_t)^2}{d-1}$, respectively. Assuming that the true variance σ_t^2 in the pen increases by coefficient $(\frac{t}{t-1})^2$ (Kristensen et al. 2012), the SSM becomes:

$$\sigma_t^2 = G_t \sigma_{t-1}^2,$$

where $G_t = (\frac{t}{t-1})^2$ for $t > 1$ ($G_1 = 1$).

Note that the SSMs satisfy the Markovian property, i.e.,

$$\Pr(X_t | \hat{X}_{t-1}, \hat{X}_{t-2}, .., \hat{X}_1) = \Pr(X_t | \hat{X}_{t-1}),$$

and because (μ_t, g_t) and σ_t are estimated independently we have

$$\Pr(X_t | \hat{X}_{t-1}) = \Pr(\ (\mu_t, g_t) \mid (\hat{\mu}_{t-1}, \hat{g}_{t-1})\ ) \cdot \Pr(\sigma_t | \hat{\sigma}_{t-1}).$$

As a result the probability distributions of these random variables can be found using Theorems 2 and 4 in Pourmoayed et al. (2016). Therefore the random information ω in the stochastic transition function $\Gamma(\cdot)$ (defined in Section 2.4) can be sampled using the conditional probability distributions of the conditional random variables (μ_t, g_t) given $(\hat{\mu}_{t-1}, \hat{g}_{t-1})$ and σ_t given $\hat{\sigma}_{t-1}$.

B Calculation of $c_k^{\texttt{cull}}(w_{ij})$ and $c_k^{\texttt{feed}}(w_{ij})$

Before calculating $c_k^{\texttt{cull}}(w_{ij})$ and $c_k^{\texttt{feed}}(w_{ij})$, we need to describe the methods of obtaining the carcass weight, leanness, feeding cost, and the settlement pork price function.

Carcass weight, leanness, feeding cost, and settlement pork price

Consider the k^{th} lightest pig with live weight w and daily growth $\breve{g}$ in the pen. The carcass weight $\tilde{w}$ can be approximated as (Andersen, Pedersen, and Ogannisian 1999)

$$\tilde{w} = 0.84w - 5.89 + e_{\texttt{c}}, \tag{21}$$

where $e_{\texttt{c}}$ is normally distributed with a mean of zero and a standard deviation of 1.96. The relation between growth rate, leanness (lean meat percentage), and feed conversion ratio varies widely between herds. Hence, these formulas must be herd specific. The leanness $\breve{w}$ can be found as (Kristensen et al.

2012)

$$\breve{w} = \frac{-30(\breve{g}-6)}{4} + 61. \tag{22}$$

The feed intake (energy intake) is modelled as the sum of feed for maintenance and feed for growth. The basic relation between daily feed intake f (FEsv[1]), live weight, and daily gain is (Jørgensen 2003)

$$f = k_1\breve{g} + k_2 w^{0.75}, \tag{23}$$

where $k_1 = 1.549$ and $k_2 = 0.044$ are constants describing the use of feed per kilograms gain and per kilograms weight, respectively. As a result the expected feed intake of the k^{th} pig over the next $\tilde{d}$ days equals

$$f_k^{\texttt{feed}}(\tilde{d}) = \mathbb{E}\left(\sum_{d=1}^{\tilde{d}} f_d\right) = \mathbb{E}\left(\sum_{d=1}^{\tilde{d}} \left(k_1\breve{g} + k_2(w+(d-1)\breve{g})^{0.75}\right)\right) = \mathbb{E}\left(\tilde{d}k_1\breve{g} + k_2\sum_{d=1}^{\tilde{d}}(w+(d-1)\breve{g})^{0.75}\right),$$

where f_d denotes the feed intake at day d calculated recursively using (23). The feeding cost for the k^{th} pig during $\tilde{d}$ days can be calculated by multiplying $f_k^{\texttt{feed}}(\tilde{d})$ to the unit feed cost $p^{\texttt{feed}}$ per FEsv

$$p_k^{\texttt{feed}}(\tilde{d}) = p^{\texttt{feed}} f_k^{\texttt{feed}}(\tilde{d}). \tag{24}$$

Consider the k^{th} lightest pig with carcass weight $\tilde{w}$ and leanness $\breve{w}$ at delivery. Under the Danish system, the settlement pork price is the sum of two linear piecewise functions related to the price of the carcass and a bonus of the leanness:

$$p_k^{\texttt{pork}}(\tilde{w},\breve{w}) = \tilde{p}(\tilde{w}) + \breve{p}(\breve{w}). \tag{25}$$

Functions $\tilde{p}(\tilde{w})$ and $\breve{p}(\breve{w})$ correspond to the unit price of carcass and the bonus of leanness for 1 kilogram meat, respectively. We define $\tilde{p}(\tilde{w})$ and $\breve{p}(\breve{w})$ based on the meat prices used in Kristensen et al. (2012) as

[1] FEsv is the energy unit used for feeding the pigs in Denmark. One FEsv is equivalent to 7.72 MJ.

$$
\tilde{p}(\tilde{w}) = \begin{cases} 0 & \tilde{w} < 50 \\ p^{\texttt{pork}} - 0.1(70-60) - 0.2(60-\tilde{w}) & 50 \leq \tilde{w} < 60 \\ p^{\texttt{pork}} - 0.1(70-\tilde{w}) & 60 \leq \tilde{w} < 70 \\ p^{\texttt{pork}} & 70 \leq \tilde{w} < 86 \\ p^{\texttt{pork}} - 0.1(\tilde{w}-86) & 86 \leq \tilde{w} < 95 \\ 9.3 & 95 \leq \tilde{w} < 100 \\ 9.1 & \tilde{w} \geq 100 \end{cases}
$$

$$
\breve{p}(\breve{w}) = 0.1(\breve{w} - 61),
$$

where $p^{\texttt{pork}}$ is the maximum price of 1 kilogram carcass (DKK/kg) when the carcass weight lies between 70 and 86 kilograms. Note that when marketing decisions are made, the culled pigs are sent to the abattoir after $t^{\mathrm{d}} = 3$ days. The additional feeding cost and reward, resulting from keeping the culled pigs in this period, can be calculated using equations (24) and (25), respectively.

Calculation of reward and feed cost

The calculations of $c_k^{\texttt{cull}}(w_{ij})$ and $c_k^{\texttt{feed}}(w_{ij})$ are rather complex due to the ordered random variables and the non-continuous function $\tilde{p}(\tilde{w})$. However, these values can be calculated using simulation with a sorting procedure as described in Algorithm 3 that takes the current weight information $w_{ij} = (\mu, \sigma, g)$ as input. First the weight of $q^{\max}$ pigs are simulated and sorted on line 3. Next, for each pig we calculate the carcass weight, leanness and the settlement pork price at slaughter (lines 6-8). Moreover, the feed cost until slaughter and for a week are found on line 9. These values are used to find $c_{k,l}^{\texttt{cull}}(w_{ij})$ and $c_{k,l}^{\texttt{feed}}(w_{ij})$ on lines 10-11. We run the simulation 1,000 times and afterwards calculate the average values for the feed cost and the reward of selling to the abattoir (lines 15-16).

Algorithm 3 - Calculation of $c_k^{\mathtt{cull}}(w_{ij})$ and $c_k^{\mathtt{feed}}(w_{ij})$

INPUT: Weight information $w_{ij} = (\mu, \sigma, g)$
FOR $l = 1$ **TO** 1000
  Draw $q^{\max}$ samples from $N(\mu, \sigma)$ and sort them in non-decreasing order $(\hat{w}_1, \ldots, \hat{w}_{q^{\max}})$;
  FOR $k = 1$ **TO** $q^{\max}$ **DO**
    Set $w := \hat{w}_k$; $w_{t^{\mathrm{d}}} = w + t^{\mathrm{d}} \cdot g/7$;
    Calculate carcass weight $\tilde{w}$ using (21) with $w = w_{t^{\mathrm{d}}}$;
    Calculate leanness $\breve{w}$ using (22) with $\breve{g} = g/7$;
    Calculate $p_k^{\mathrm{pork}}(\tilde{w}, \breve{w})$ using (25);
    Calculate feed cost $f_k^{\mathrm{feed}}(\tilde{d})$ for $\tilde{d} = 7$ and $\tilde{d} = t^{\mathrm{d}}$ using (24) with w and $g/7$;
    Calculate $c_{k,l}^{\mathtt{cull}}(w_{ij}) = p_k^{\mathrm{pork}}(\tilde{w}, \breve{w}) - f_k^{\mathrm{feed}}(t^{\mathrm{d}})$;
    Set $c_{k,l}^{\mathtt{feed}}(w_{ij}) = f_k^{\mathrm{feed}}(7)$;
  END FOR
END FOR
FOR $k = 1$ **TO** $q^{\max}$ **DO**
  $c_k^{\mathtt{cull}}(w_{ij}) = \sum_l c_{k,l}^{\mathtt{cull}}(w_{ij})/1000$;
  $c_k^{\mathtt{feed}}(w_{ij}) = \sum_l c_{k,l}^{\mathtt{feed}}(w_{ij})/1000$;
END FOR